

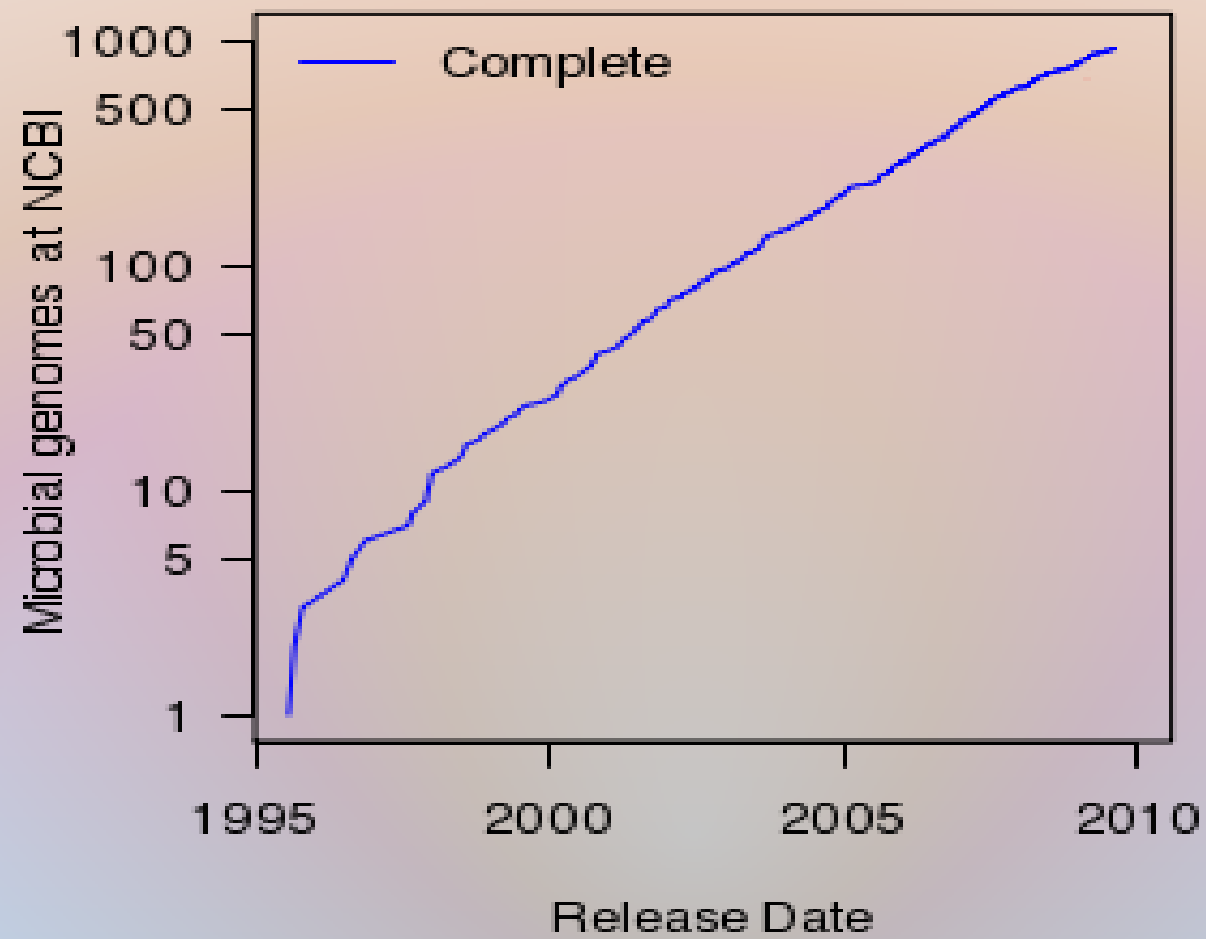
Assigning protein functions by comparative genome analysis using protein phylogenetic profiles

Content

- Introduction
- Protein phylogenetic profiles
- Proceedings
- Todays data situation

Introduction

- Rising genomic sequence data

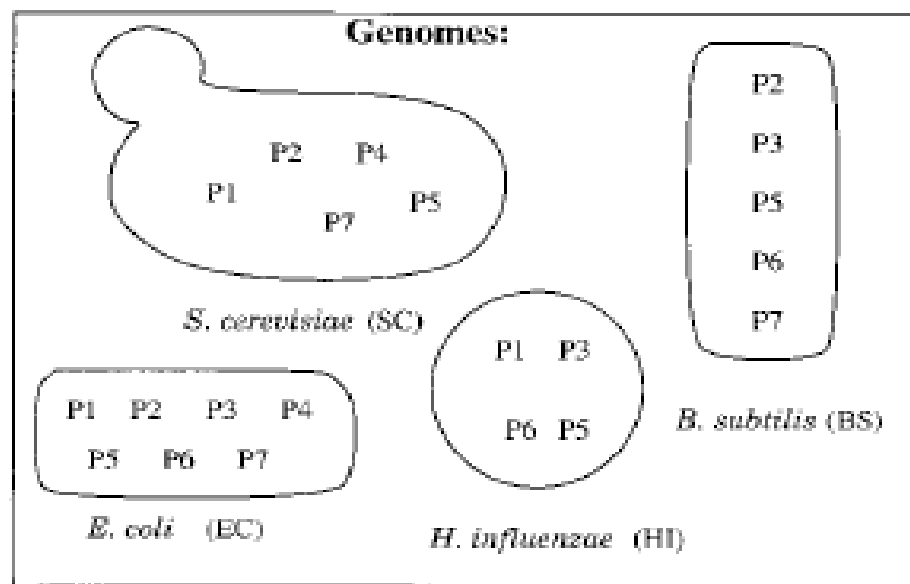


Introduction

- Rising genomic sequence data
- Assign protein functions from genomic sequence
- Homology modeling is used for detecting functional related proteins in different organisms

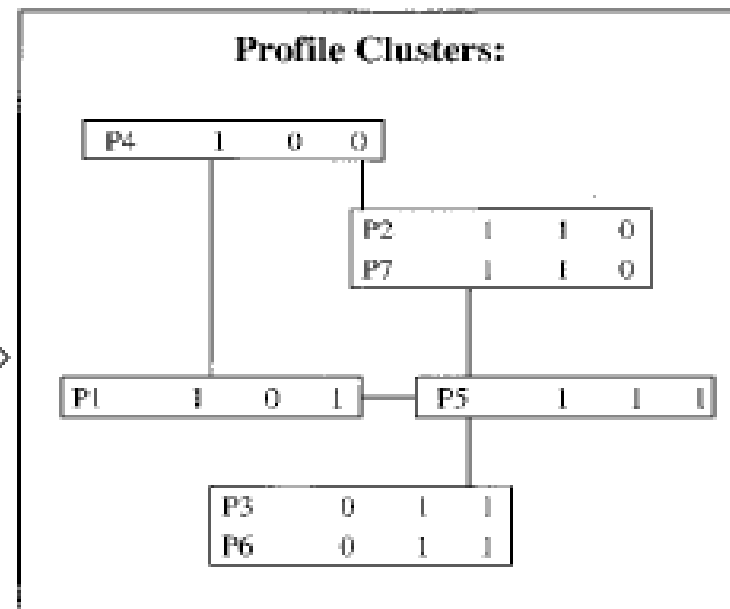
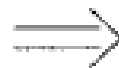
Protein Phylogenetic Profiles

- Correlated evolution
 - Elimination/Preserving of functional related proteins during evolution
- Should detect proteins that are functionally linked
 - Homologs in the same subset of organisms
 - e.g. flagellar proteins
- No sequence similarity between sequences
 - => sequence alignment techniques are useless



Phylogenetic Profile:

	EC	SC	BS	HI
P1	1	1	0	1
P2	1	1	1	0
P3	0	1	1	1
P4	1	0	0	0
P5	1	1	1	1
P6	0	1	1	1
P7	1	1	1	0

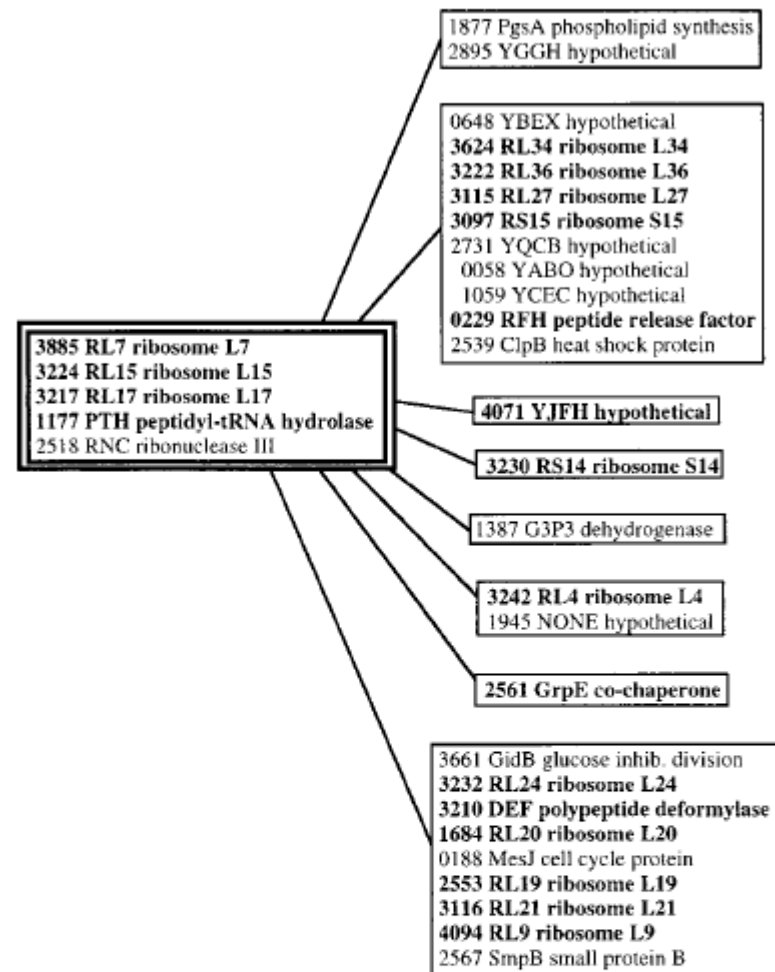


Conclusion: P2 and P7 are functionally linked,
P3 and P6 are functionally linked

A

Initial Profile

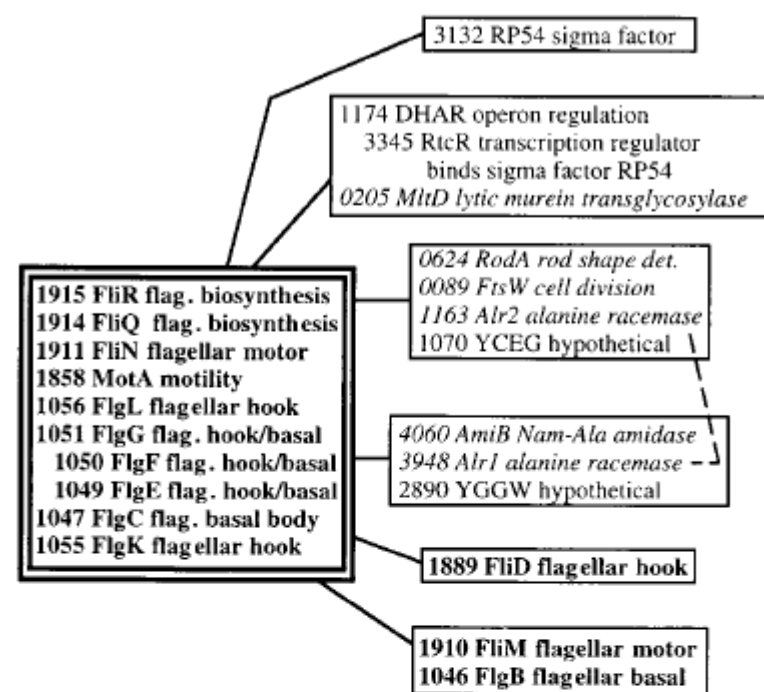
One bit different



B

Initial Profile

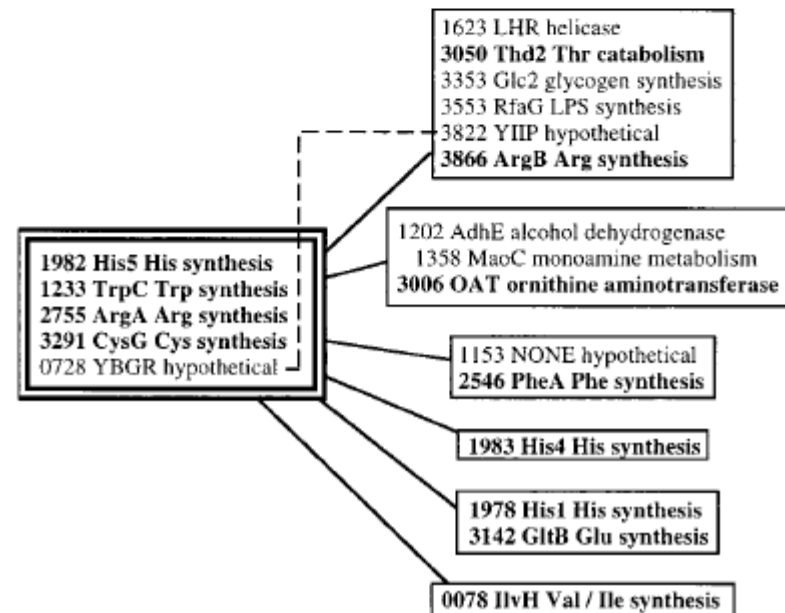
One bit different



C

Initial Profile

One bit different



What is the Handicap to this point?

- Only 16 genomes (1999)
- Ignores phylogenetic relationships between organisms

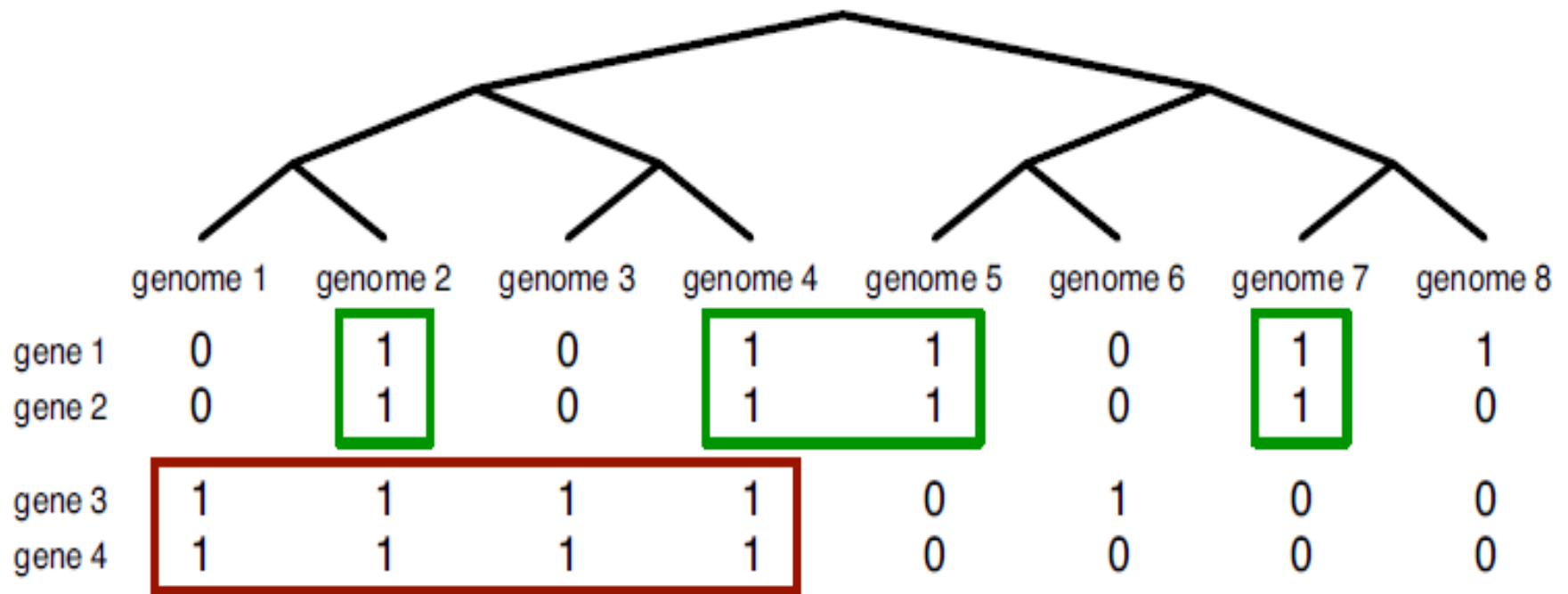
Proceedings

An improved method for identifying functional linked proteins using phylogenetic profiles (2007)

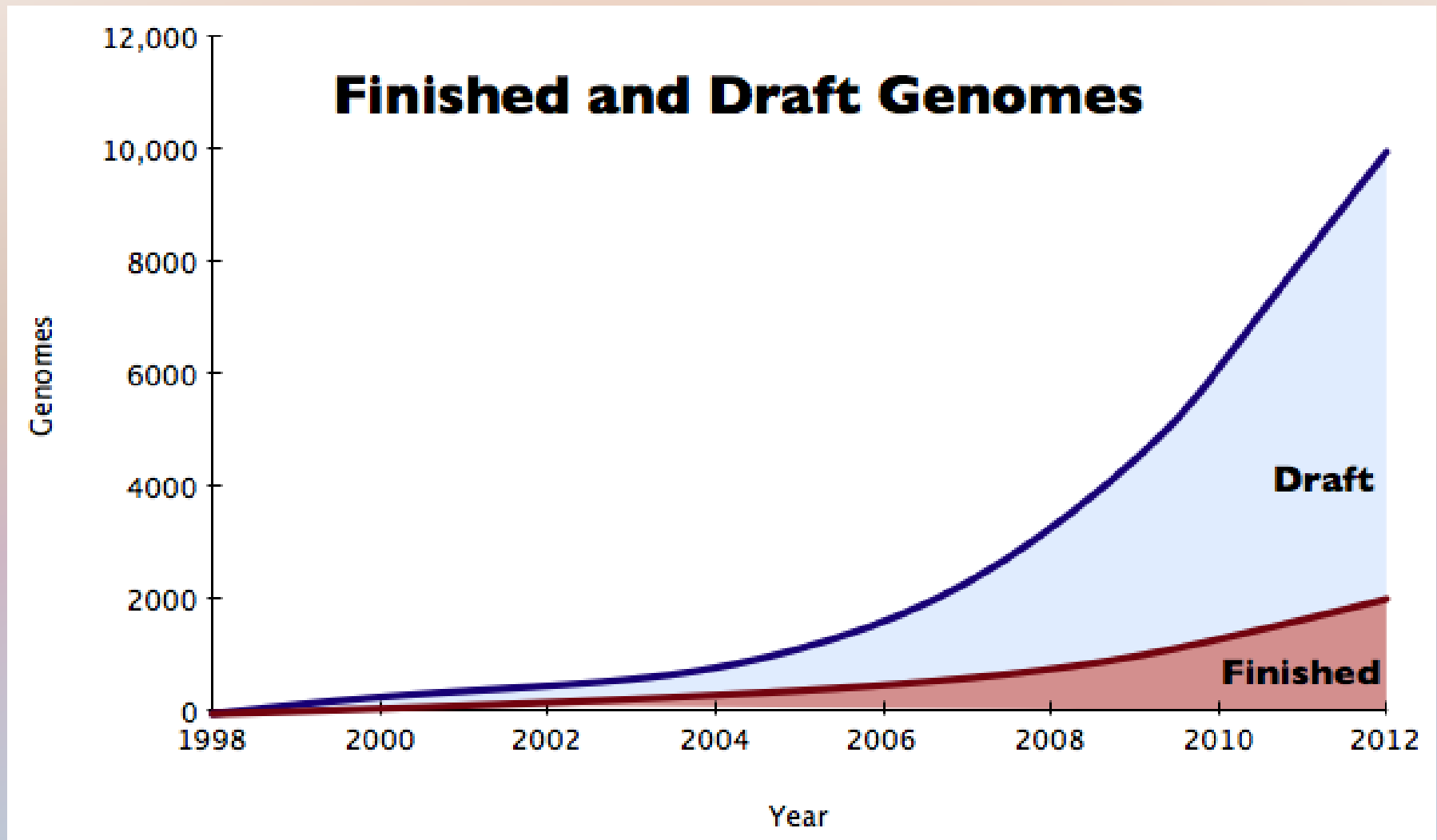
What is the data situation at this point?

- ~400 Genomes available
- Still using BLAST for Homology Detection
- Used only 214 genomes for the profile vector leaving out similar genomes

Phylogenetic relationships



Today's data situation



Today's data situation

- >1100 Genomes available
- Advanced homology detection methods

Any questions?

References

- Matteo Pellegrini, Edward M. Marcotte, Michael J. Thompson, David Eisenberg and Todd O. Yeates *Assigning protein functions by comparative genome analysis: Protein phylogenetic profiles*. Proc Natl Acad Sci U S A. 1999 Apr 13;96(8):4285-8.
- Cokus S, Mizutani S, Pellegrini M. *An improved method for identifying functionally linked proteins using phylogenetic profiles*. BMC Bioinformatics. 2007 May 22;8 Suppl 4:S7.
- www.wikipedia.org
- <http://www.ncbi.nlm.nih.gov/>